Package 'nis'

October 19, 2011

Version 0.1-1

Date 2011-10-18

Title Nationwide Inpatient Sample tools for R

Authors@R person(given = c(''Jeremy'', ''Thoms''), family = ''Thoms'', email = ''jthetzel@gmail.com'')

Author Jeremy T. Hetzel <jthetzel@gmail.com>

Maintainer Jeremy T. Hetzel <jthetzel@gmail.com>

Depends R (>= 2.12.0)

Description The nis package provides tools for working with data from the Nationwide Inpatient Sample. Included is a convenience function to generate SQL code for importing the NIS flat files into a MySQL database.

License GPL-3

URL http://code.google.com/p/nis

BugReports http://code.google.com/p/nis

LazyLoad yes

LazyData yes

Collate 'generateSQL.R'

R topics documented:

---

nis-package   *Tools for working with data from the Nationwide Inpatient Sample*

Description

Tools for working with data from the Nationwide Inpatient Sample

1

---

generateSQL   *Generate SQL code to convert NIS ascii flat files to SQL database*

Usage

```
generateSQL(years, files, type, remove.capitalization =
T, db.table = NULL, layouts.uri = NULL, old = NULL, new =
NULL, normalize.dx = F, normalize.pr = F, names = NULL)
```

Arguments

years	A vector containing the years of data to include. The order of `years` must be the same as file locations provided in the `files` parameter. Currently, only years 1998 through 2009 are supported.
files	A vector containing the file locations of the ascii flat files. The order of `files` must be the same as the years provided by the `years` parameter.
type	The type of NIS data being loaded. Acceptable values are `"core"`, `"hospitals"`, `"severity"`, and `"groups"`, for the NIS Core, Hospitals, Severity, and Dx Pr Groups data files, respectively.
remove.capitalization	
	optional A logical indicating whether the variable names should be converted to all lower case. Default is `TRUE`.
db.table	optional A character indicating the desired database table name. If not specified, will default to the `type` parameter (i.e. core, hospitals, severity, or groups)
layouts.uri	optional A list specifying the uris to use to fetch the layouts of the ascii flat files. This should usually not be specified, in which case the default uris are used, as listed in details below.
old	optional A vector of variable names to be replaced by `new`. Must be in the same order as the `new` parameter.
new	optional A vector of new variable names to replace those specified in `old`. Must be in the same order as the `old` parameter.
normalize.dx	Logical indicting whether to provide additional SQL statements to normalize the diagnosis (dx1, dx2, etc.) variables into a sperate table. Default is `FALSE`.
normalize.pr	Logical indicting whether to provide additional SQL statements to normalize the procedure (pr1, pr2, etc.) variables into a sperate table. Default is `FALSE`.
names	Character vector giving names for the three columns of the normalized diagnosis and procedure tables. Only applicable if `normalize.dx` or `normalize.pr` are `TRUE`. Default is NULL, in which case the columns are named "keyid", "icd9", and "variable", for the record number, ICD9 code, and variable name (i.e. dx1 or dx2), respectively.

Value

A list of two character vectors:

createTable	A SQL statement to create an empty table with with variables appropriate for NIS data.

loadData   One or more SQL statements to load data from the specified `files` into a MySQL table using the `load data infile` SQL statement.

The `generateSQL` function is a convenience function that returns SQL code to be used for creating appropriate tables for a MySQL database and loading data from the flat ascii files provided by HCUP.

Details Unless `layouts.uri` is specified, `generateSQL` uses the parses layout information from the NIS website to determine the appropriate variable names and fixed-width locations of data in the NIS ascii flat files. By default, the `layouts.uri` object is the following list: `layouts.uri <-`
`list( "1998" = 'http://www.hcup-us.ahrq.gov/db/nation/nis/tools/stats/NIS_1998_COREv2.TXT',`
`"1999" = 'http://www.hcup-us.ahrq.gov/db/nation/nis/tools/stats/NIS_1999_COREv2.TXT',`
`"2000" = 'http://www.hcup-us.ahrq.gov/db/nation/nis/tools/stats/NIS_2000_CORE.TXT',`
`"2001" = 'http://www.hcup-us.ahrq.gov/db/nation/nis/tools/stats/NIS_2001_CORE.TXT',`
`"2002" = 'http://www.hcup-us.ahrq.gov/db/nation/nis/tools/stats/FileSpecifications_NIS_2002_CORE`
`"2003" = 'http://www.hcup-us.ahrq.gov/db/nation/nis/tools/stats/FileSpecifications_NIS_2003_CORE`
`"2004" = 'http://www.hcup-us.ahrq.gov/db/nation/nis/tools/stats/FileSpecifications_NIS_2004_CORE`
`"2005" = 'http://www.hcup-us.ahrq.gov/db/nation/nis/tools/stats/FileSpecifications_NIS_2005_Core`
`"2006" = 'http://www.hcup-us.ahrq.gov/db/nation/nis/tools/stats/FileSpecifications_NIS_2006_Core`
`"2007" = 'http://www.hcup-us.ahrq.gov/db/nation/nis/tools/stats/FileSpecifications_NIS_2007_Core`
`"2008" = 'http://www.hcup-us.ahrq.gov/db/nation/nis/tools/stats/FileSpecifications_NIS_2008_Core`
`"2009" = 'http://www.hcup-us.ahrq.gov/db/nation/nis/tools/stats/FileSpecifications_NIS_2009_Core`
`)`

Create SQL statement to upload NIS Core data from 1998-2009 to MySQL database # Years to include years <- c(seq(1998, 2009))

Location of the NIS ascii flat files files <- c('/data/nis/NIS_1998_Core.ASC', '/data/nis/NIS_1999_Core.ASC', '/data/nis/NIS_2000_Core.ASC', '/data/nis/NIS_2001_Core.ASC', '/data/nis/NIS_2002_Core.ASC', '/data/nis/NIS_2003_Core.ASC', '/data/nis/NIS_2004_Core.ASC', '/data/nis/NIS_2005_Core.ASC', '/data/nis/NIS_2006_Core.ASC', '/data/nis/NIS_2007_Core.ASC', '/data/nis/NIS_2008_Core.ASC', '/data/nis/NIS_2009_Core.ASC')

Generate SQL statemebts sql.core <- generateSQL(years = years, files = files, type = "core")

Output SQL statements to a file #cat(sql.core$createTable, file = "makeTableCore.sql") #cat(sql.core$uploadData, file = "uploadDataCore.sql")

http://www.hcup-us.ahrq.gov/nisoverview.jsp

Index